All Included.

目录

1

1. A program finishes a task in 10 hours and 70% of this program can be parallelized. According to the Amdahl's law, how much will it cost if the program uses 4 processors to work on this task?

A. 4.75 h
B. 7 h
C. 4 h
D. 3.57 h

Solution:

$$10 \times \left(1 - 70\% + \frac{70\%}{4}\right) = 4.75$$

2. If the L1 cache has 20% miss rate with no penalty on hits, while the L2 cache has a 10% miss rate with 10ns hit time, i.e., to transfer data to L1 cache. The system memory

has an access latency of 70ns, and it has 1% possibility that it will access the virtual memory space on disk, which introduces 1ms penalty. What is the average memory access time for a load word operation?

A. 203.4 ns
B. 212.4 ns
C. 2012.4 ns
D. 3.6 ns

Solution:

$$L1\ time = 0 + 20\% \times miss\ penalty_{L1}$$
$$miss\ penalty_{L1} = 10 + 10\% \times miss\ penalty_{L2}$$
$$miss\ penalty_{L2} = 70 + 1\% \times 1000000$$
$$L1\ time = 203.4\ ns$$

3. Consider the MIPS program:

i) DIV.D F4, F2, F1
ii) SUB.D F5, F4, F1
iii) ADD.D F4, F8, F2
iv)L.D F2, 8(R1)
(Hint: op dst, src1, src2)

We always want to exploit ILP as much as possible, so we often reschedule the code and implement out of order execution to minimize the stalls, but we may encounter data hazards. In our case here, which of the following statements is correct?
A. WAR: ii) and iii)
B. WAR: i) and ii)
C. WAW: iii) and iv)
D. RAW: i) and iv)

2

Suppose that there is a processor and the memory hierarchy only contains a cache and main memory, and it applies virtual memory with a TLB and a two-level page table. Suppose the TLB miss rate is 5% and cache miss rate is 10%. What's the average main memory access times in one memory reference instruction?

* 0.110
* 0.105
* 0.100

Reason:

* average memory access times per store/load instruction: 5%*2(TLB miss -> 2 times page table access) + 1 = 1.1
* main memory: 1.1*10%=0.11

廖伟鑫

```
### Computer architecture Hw4
* Q: Please choose the correct statements regarding the single-cycle processor introduced in the lecture.
    * Load and store instructions use ALU, Sign Extend and Data Memory.
    * Jump instructions use Instruction Memory, ALU and PC.
    * R-type instructions use Register File, PC and Sign Extend.

* Ans: A
    * A is correct.
    * Jump instructions don't use the ALU part.
    * R-type instructions have no immediate operation, don't use Sign Extend.
```

Problems in this exercise refer to the following fragment of MIPS code:

```
sw r16,12(r6)
lw r16,8(r6)
beq r5,r4,Label     # Assume r5!=r4
add r5,r1,r4
slt r5,r15,r4
```

Assuming stall-on-branch and no delay slots, what speedup is achieved on this code if branch outcomes are determined in the ID stage, relative to the execution where branch outcomes are determined in the EX stage?

A:1.10x

B:1.07x

C:1.13x

Instruction	Pipeline Stage	Cycles

SW R16,12(R6)
LW R16,8(R6)
BEQ R5,R4,Lbl
ADD R5,R1,R4
SLT R5,R15,R4
	IF ID EX MEM WB
	IF ED EX MEM WB
	IF ID EX MEM WB
	*** *** IF ID EX MEM WB
	IF ID EX MEM WB
		11

4.10.3 Stall-on-branch delays the fetch of the next instruction until the branch is executed. When branches execute in the EXE stage, each branch causes two stall cycles. When branches execute in the ID stage, each branch only causes one stall cycle. Without branch stalls (e.g., with perfect branch prediction) there are no stalls, and the execution time is 4 plus the number of executed instructions. We have:

Instructions Executed	Branches Executed	Cycles with branch in EXE	Cycles with branch in ID	Speedup
5	1	4 + 5 + 1*2 = 11	4 + 5 + 1*1 = 10	11/10 = 1.10

3

Which of the following instructions contains load-use data hazard ?

-add $s1, $t1, $t3
 sub $t2, $s1, $t0
 lw $t1, 0($t0)

-add $t3, $t1, $t2
 sw $t3, 12($t0)
 lw $t4, 8($t0)

-beq $t1, $t3, 28
 and $t2, $t4, $t5
 or $t8, $t6, $t2

Consider 5 memory banks used for vector loads and stores, where each bank i ($0 \leqslant i < 5$) contains the bytes with byte addresses X where X mod 5 = i (note 5 is the number of memory banks).

Suppose the banks are not pipelined and a bank's busy time is 6 cycles and memory access latency is 15 cycles. Consider an array with 7 elements and each element is 1-byte.

Specify the number of clock cycles that are required to load all the elements in the array.

Assume the address generator can generate only 1 (byte) address per clock cycle and each memory bank has independent bank addressing

-22

-23

-21

Suppose the initial value in register $1 is 202
 xor $1, $1, $1
what is the value in register $1 after executing the instruction below
- 0
- 202
- 404

4

Q1.(Amdahl's Law)A task can be divided into three parts:P1,P2,P3. They respectivly take 20%,30%,50% of the time. Now we implement a 2x speedup to P1 and a 3x speed up to P2 .What's the overall speedup?
- 1.43
- 0.7
- 1.67

Q2. For MIPS-32 ISA,what type of instrution is load instruction?
- I-type
- R-type
- J-typr

Q3.Using the following instruction mix, what is the CPI, assuming that each state in the multicycle CPU requires 1 clock cycle?

load	25%
store	10%
barnches	10%
jumps	5%
ALU	50%

- 4.1
- 4.0
- 4.2

5

Corresponding concepts:: cache
You are experimenting with a computer having an L1 data cache and a main memory (you

exclusively focus on data accesses). The latencies (in CPU cycles) of the different kinds of accesses are as follows: cache hit, 1 cycle; cache miss, 95 cycles; main memory access with cache disabled, 90 cycles. When you run a program with an overall miss rate of 10%, what will the average memory access time (in CPU cycles) be?
-10.4
-9.9
-6.4

Suppose the initial value in register $1 is 202

```
xor $1, $1, $1
```

what is the value in register $1 after executing the instruction below
- 0
- 202
- 404

Corresponding concepts: **Processor-Pipeline-Branch Prediction**

What is the accuracy of always-taken and always-not-taken predictors for the following repeating pattern (e.g., in a loop) of branch outcomes: T, NT, T, T, NT
-60%, 40%
-40%, 60%
-60%, 60%

6

In the process of main-cache address transformation, the lowest block collision rate is ().

A. Fully associated mapping

B. Direct mapping

C. Set associated mapping

D. None of the above

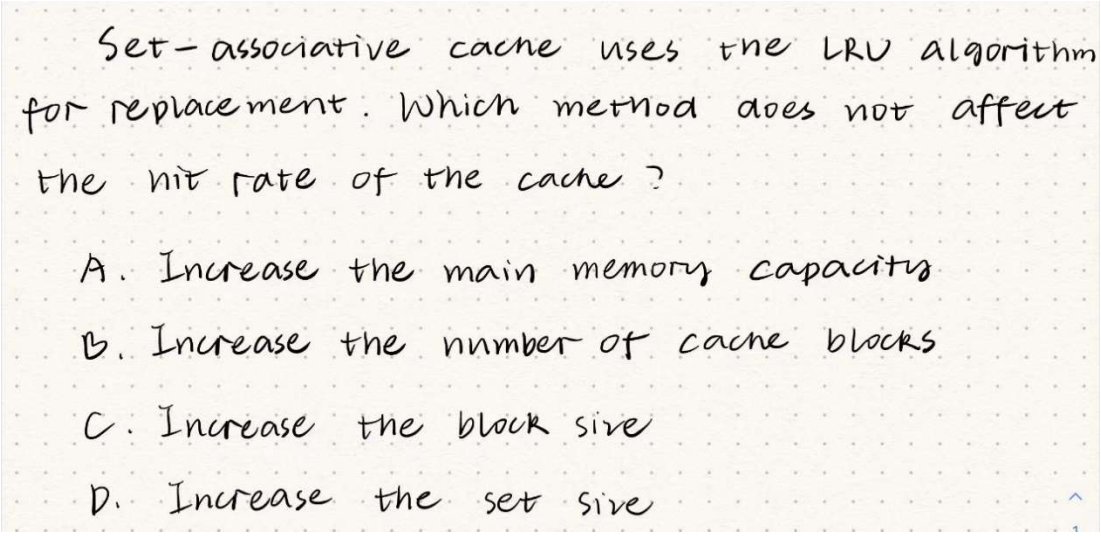
According to the level of pipeline, the pipeline can be divided into 3 levels, they are not included

A. Component level pipeline

B. Processor level pipeline

C. Between processors pipeline

D. Registers level pipeline



7

题目 10

There is a program that takes 10 seconds to run on computer A with a clock frequency of 2GHz. Now we try to help computer designers build computer B that runs the program in 8 seconds. The designers have determined that the clock rate can be significantly increased, but this may affect the design of the rest of the CPU, so that computer B needs 1.2 times as many clock cycles to run the program as Computer A. What should we recommend the computer designers set as the design goal of the constant frequency?

a) 3GHz

b) 4GHz

c) 5GHz

Q: Given the instructions in the Pipelined MIPS processor as below, please judge these instructions will cause hazard in which way?(We don't consider the optimization about the hazards)
add $1,$2,$3
add $4,$1,$3

- Data hazard
- Structure hazard
- Control hazard

Q: Which of the following operations cannot be done in a MIPS instruction set with a single instruction?

A. Add the value in memory **m** to the value in register **r1** and save the result in **r1**.
B. Add the value in register **r2** to the value in register **r1** and save the result in **r3**.
C. Save the value in memory **m** in register **r1**.

8

(Cache)

Assume that main memory accesses take 68 ns. The following Table 1 data for L1 caches attached to each of two processors, P1 and P2.

	L1 Size	L1 Miss Rate	L1 Hit Time
P1	2KB	13.4%	0.72ns
P2	4KB	7.8%	0.87ns

What is the AMAT (Average Memory Access Time) for P1 and P2?

a) 9.832ns, 6.174ns

b) 9.736ns, 6.106ns

c) 9.986ns, 6.306ns

题目 11
Assume now that we have a non-pipelined processor, which has a cycle time of 10ns.
By pipelining, the latencies of 5 pipeline stages are as following.
And each stage also adds 20ps due to register setup delay.

What is the best speedup we can get compared to the original processor?

stage	IF	ID	EX	MEM	WB
latency / ns	1	1.5	4	3	0.5

a) 2.488
b) 1.010
c) 2.010

题目 12
Q: Given the instructions in the Pipelined MIPS processor as below, please judge these instructions will cause hazard in which way? (We don’t consider the optimization of the hazards)
add $1,$2,$3
add $4,$1,$3

- Data hazard
- Structure hazard
- Control hazard

9

Discuss the efficiency of the five-stage assembly line, ignoring the risk, which of the following assertions is true?

①*Allowing some instructions to use fewer segments does not improve performance because throughput is determined by the clock cycle. The number of pipeline sections required for each instruction only affects the delay time, but not the throughput of the pipeline*

②*It is not possible to reduce the number of clock cycles required for ALU instructions because they need to write back the results. However, branch and jump instructions can reduce the number of clock cycles, so there are opportunities for performance improvement*

③*Jump, branch, and ALU instructions using fewer segments than five (the number of segments used by the LOAD instruction) will increase pipeline performance in all cases*

④*Instead of trying to reduce the number of clock cycles required for an instruction, we can extend the number of lines. Although each instruction takes more clock cycles, the length of the clock cycle is shorter, which increases instruction throughput and thus the overall performance of program execution*

A. ①④
B. ①③
C. ②④

TLB works on which layer of memory hierarchy?

A. main memory – disk
B. cache – main memory
C. register – memory

Here's a 5-stage pipelined processor. Suppose the latency of each stage is:

IF stage	ID stage	EX stage	MEM stage	WB stage
3ns	2.5ns	2ns	3.5ns	6ns

What is the clock cycle time of the 5-stage pipelined processor?

A. 6ns
B. 2ns
C. 3.5ns

10

Q: Which of the following statement about memory is true?

- The latency of cache is usually shorter than main memory.
- Both SRAM and DRAM are non-volatile memory.
- Flash memory is slower than disk.

***Q*: You are given a non-pipelined processor design which has a cycle time of 10ns and average CPI of 1.4. Which one is NOT correct in the following statements?**

A. If the pipeline stalls 20% of the time for 1 cycle and 5% of the time for 2 cycles, the new CPI is 1.8

B. The best speed up we could get by pipelining it into 5 stages is 5X

C. If the 5 stages are 1ns, 1.5ns, 4ns, 3ns, and 0.5ns, the best speedup is 2.5X

1) The address in the main memory is known as -

1. Physical address
2. Logical address
3. Memory address

2) CISC stands for -

1. Complex Instruction Set Computer
2. Complete Instruction Sequential Compilation
3. Complex Instruction Sequential Compiler

3) RISC stands for -

1. Reduce Instruction Set Computer
2. Risk Instruction Sequential Compilation
3. Risk Instruction Source Compiler

11

1.Consider 8 memory banks used for vector loads and stores, where each bank i ($0 \leqslant i < 8$) contains the bytes with byte addresses X where X mod 8 = i.
Suppose the banks are not pipelined and a bank's busy time is 6 cycles and memory access latency is 12 cycles. Consider an array with 64 elements and each element is 1-byte.
Assume the address generator can generate only 1 (byte) address per clock cycle and each memory bank has independent bank addressing.
Please specify the number of clock cycles that are required to load all the elements in the array with a stride of 4.

A- 200
B- 76
C- 768

2.Which one is true in memory hierarchy?
A.What is in L1 is a subset of what is in L2
B.What is in L1 has nothing to do with what is in L2
C.L1 is larger and closer to processor than L2

3. Suppose you have two execution pipelines, each capable of beginning execution of one instruction per cycle, and enough fetch/decode bandwidth in the front end so that it will not stall your execution. Assume results can be immediately forwarded from one execution unit to another, or to itself. Further assume that the only reason an execution pipeline would stall is to observe a true data dependency. The following codes is a loop and the latency for various operation beyond single cycle is also listed. How many cycles does the loop require ?

Loop: LD F2,0(RX)
IO: DIVD F8,F2,F0

I1: MULTD F2,F6,F2
I2: LD F4,0(Ry)
I3: ADDD F4,F0,F4
I4: ADDD F10,F8,F2
I5: ADDI Rx,Rx,#8
I6: ADDI Ry,Ry,#8
I7: SD F4,0(Ry)
I8: SUB R20,R4,Rx
I9: BNZ R20,Loop

(latency beyond single cycle: Memory LD :+4 Memory SD :+1 Integer ADD, SUB :+0 Branches :+1 ADDD :+1 MULTD :+5 DIVD :+12)
A.22
B.20
C.2147483647

12

题目 7
Cache is a small-capacity memory between CPU and (), which can provide CPU with high speed Instructions and Data
-main memory
-Register
-hard disk

题目 8
The design idea of von Neumann type computer is
-Store the program and execute it in address order
-Store data and execute in address order
-Store the program and execute it out of order

题目 9
In the hierarchical structure of the computer system, the hardware level is
-Machine language level
-Assembly language level
-Advanced language level

13

Suppose there are 1000 memory accesses for a cache. 400 remains after L1 cache and 100 remains after L2. Q: The Miss Rate of L1 cache and ***Global*** Miss Rate of L2 cache.
a)0.4,0.1
b)0.4,0.25
c)0.6,0.25

If fraction F of a program is enhanced by a factor of E when executed on an enhanced machine, then derive the expression of speedup over the original machine. Based on that expression, what is the speedup? And if the old execution time of a program is T, what about the new execution time?

a)$\frac{1}{(1-F)+\frac{F}{E}}$;T[(1-F)+$\frac{F}{E}$]

b) $(1-F)+\frac{F}{E}$; T[(1-F)+$\frac{F}{E}$]

c) $\frac{1}{(1-F)+\frac{F}{E}}$; $\frac{T}{(1-F)+\frac{F}{E}}$

For a typical cache, suppose the index is 8 bits and the block offset is 6 bits. If the cache is an 4-way associative cache, and the last 16 bits of load data address is 1010010001111100, then what's the size of the data?
A. Unknown.
B. 4 bytes
C. 64 bytes

14

Which of the following is the advantage of load-store instruction category?
A. The clock cycles of each instruction are near.
B. It can directly access the number of operations of registers.
C. The quantity of its objective code is small.

Which of the following in a cache address, specifies the exact location in the cache line where the requested data exists?
A.Block Offset
B.Block Number

C.Tag
D.Block Index

Which of the following is a disadvantage of Pipelining?
A. Both B and C
B. The design of pipelined processor is complex and costly to manufacture.
C. The instruction latency is more.
D. Cycle time of the processor is reduced.